



# SEQUENCE LISTING

<110> KIM, Young Tae  
LEE, Jae Hyung

<120> Gene involved in the biosynthesis of carotenoid and marine microorganism, paracoccus haeundaesis, producing the carotenoid

<130> 428.1056

<140> US 10/551,508  
<141> 2005-09-29

<150> PCT/KR2004/000752  
<151> 2003-03-31

<150> KR2003-20222  
<151> 2003-03-31

<150> KR2003-20023  
<151> 2003-03-31

<160> 18

<170> KopatentIn 1.71

<210> 1  
<211> 21  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> forward primer for Paracoccus haeundaesis 16S rDNA

<400> 1  
cataagtaat tatggttttg t 21

<210> 2  
<211> 18  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> reverse primer for Paracoccus haeundaesis 16S rDNA

<400> 2  
cgcttcctta gaaaggag 18

<210> 3  
<211> 1454  
<212> DNA

<213> Paracoccus haeundaesis

<400> 3

```
caacttgaga gtttgatcct ggctcagaac gaacgctggc ggcaggctta acacatgcaa      60
gtcgagcgag accttcgggt ctagcggcgg acgggtgagt aacgcgtggg aacgtgccct      120
tctctacgga atagccccgg gaaactggga gtaataccgt atacgccctt tgggggaaag      180
atttatcgga gaaggatcgg cccgcgttgg attaggtagt tgggtgggta atggcccacc      240
aagccgacga tccatagctg gtttgagagg atgatcagcc aactggggac tgagacacgg      300
cccagactcc tacgggaggg agcagtgggg aatcttagac aatgggggca accctgatct      360
agccatgccg cgtgagtgat gaaggcctta ggggtgtaaa gctctttcag ctgggaagat      420
aatgacggta ccagcagaag aagccccggc taactccgtg ccagcagccg cggtaatatcg      480
gagggggcta gcgttggtcg gaattactgg gcgtaaagcg cacgtgggcg gactggaaag      540
tcagaggtga aatcccaggg ctcaaccttg gaactgcctt tgaaactatc agtctggagt      600
tcgagagagg tgagtggaat tccgagtgtg gaggtgaaat tcgtagatat tcggaggaac      660
accagtggcg aaggcggctc actggctcga tactgacgct gaggtgcaa agcgtgggga      720
gcaaacagga ttagataccc tggtagtcca cgccgtaaac gatgaatgcc agacgtcggc      780
aagcatgctt gtcggtgtca cacctaacgg attaagcatt ccgcctgggg agtacggtcg      840
caagattaaa actcaaagga attgacgggg gcccgacaaa gcggtggagc atgtggttta      900
attcgaagca acgcgcagaa cttaccaac cttgacatg gcaggaccgc tggagagatt      960
cagctttctc gtaagagacc tgcacacagg tgctgcatgg ctgtcgtcag ctcgtgtcgt     1020
gagatgttcg gttaagtccg gcaacgagcg caaccacgt ccctagttgc cagcattcag     1080
ttgggcactc tatggaaact gccgatgata agtcggagga aggtgtggat gacgtcaagt     1140
tctcatggcc cttacgggtt gggctacaca cgtgctacaa tggtggtgac agtgggttaa     1200
tccccaaaag ccatctcagt tcggattgtc ctctgcaact cgagggcatg aagttggaat     1260
cgctagtaat cgcggaacag catgccgcgg tgaatacgtt cccgggcctt gtacacaccg     1320
cccgtcacac catgggagtt ggttctaccc gacgacgctg cgctaacctt cggggggcag     1380
gcggccacgg taggatcagc gactgggggtg aagtcgtaac aaggtagccg taggggaacc     1440
tgcggctgga tcac                                             1454
```

<210> 4

<211> 6223

<212> DNA

<213> crt gene

<400> 4

```
gttccacgac tggggcatcc ccacgaccgc gtcgctgcgc gccatcgcg cgatgatggg 60
gccggaccgg gttctggtcg ggtcgggagg ggtgcgtcac gggctggacg ccgcgcgggc 120
catccgcctc ggcgcggacc tcgtggggca ggcggcccgc gcgctgcccg ccgcgcgcca 180
cagcgccgag gccctgtccg atcacctgtc cgacgtcgtg acccagctgc gcatcgcgat 240
gttctgcacc ggatcggggc accttgacgc gctgcgctgc gcgcctctgc tgggtgccggg 300
gccgggtggc caatggtcgc aagcaacggg gatggaaacc ggcgatgcgg gactgtagtc 360
tgcgcggatc gccgggtccg gggacaagat gagcgacat gccctgccc aggcagatct 420
gaccgccacc agcctgatcg tctcggggcg catcatcgcc gcgtggctgg ccctgcatgt 480
gcatgcgctg tggtttctgg acgcggcggc gcatcccatc ctggcgatcg cgaatttcct 540
ggggctgacc tggctgtcgg tcggtctgtt cttcatcgcg catgacgca tgcacgggtc 600
ggtcgtgccg gggcgccgc gcggcaatgc ggcgatgggc cagctggtcc tgtggctgta 660
tgccggattt tcgtggcgca agatgatcgt caagcacatg gccatcacc gccataccgg 720
aaccgacgac gaccccgatt tcgaccatgg cggcccggtc cgctggtacg cgcgcttcat 780
cggcacctat ttcggctggc gcgaggggct gctgctgcc gtcacgtga cggctctatgc 840
gctgatcctg ggggatcgtt ggatgtacgt ggtcttctgg ccgctgccgt cgatcctggc 900
gtcgatccag ctgttcgtgt tcggcacctg gctgccgcac cggcccgccc acgacgcgtt 960
cccggaccgc cataatgcgc ggtcgtcgcg gatcagcgac cccgtgctgc tgetgacctg 1020
ctttcacttt ggtggttatc atcacgaaca ccacctgcac ccgacggtgc cttggtggcg 1080
cctgcccagc acccgacca agggggacac cgcacgacca atttcctgat cgtcgtcgcc 1140
accgtgctgg tgatggagtt gacggcctat tccgtccacc gttggatcat gcacggcccc 1200
ctgggctggg gctggcacia gtcccaccac gaggaacacg accacgcgct ggaaaagaac 1260
gacctgtacg gcctgggtct tgcggtgatc gccacggtgc tgttcacggt gggctggatc 1320
tgggcgcggg tcctgtggtg gatcgctttg ggcacgaccg tctatgggct gatctatttc 1380
gtcctgcacg acgggctggt tcatcagcgc tggccgttcc gctatatccc gcgcaagggc 1440
tatgcccgcc gcctgtatca ggcccaccgc ctgcaccacg cggctgaggg acgcgaccat 1500
tgcgtcagct tcggcttcat ctatgcgccg ccggtcgaca agctgaagca ggacctgaag 1560
acgtcggggc tgctgcgggc cgaggcgag gagcgcacgt gacccatgac gtgctgctgg 1620
```

caggggcggg ccttgcaaac gggctgatcg ccctggcgct gcgcgcggcg cggcccgcacc 1680  
tgcggggtgct gctgctggat catgcggcgg gaccgtcaga cggccataacc tggctcctgcc 1740  
acgacccccga tctgtcgccg cactggctgg cgcggtgaa gccctgcgc cgcgccaact 1800  
ggcccgcacca ggaggtgcgc tttccccgcc atgcccggcg gctggccacc ggttacgggt 1860  
cgctggacgg ggcggcgctg gcggatgcgg tggcccggtc gggcgccgag atccgctgga 1920  
acagcgacat cgccctgctg gatgaacagg gggcgacgct gtcctgcggc acccggatcg 1980  
aggcgggcgc ggtcctggac gggcgcgggc cgcagccgtc gcggcatctg accgtggggt 2040  
tccagaaatt cgtgggcgtc gagatcgaga ccgactgcc ccacggcgtg cccgcccga 2100  
tgatcatgga cgcgaccgtc acccagcagg acgggtaccg attcatctat ctgctgccct 2160  
tctctccgac gcgcctcctg atcgaggaca ctgcctatc cgatggcggc aatctggacg 2220  
acgacgcgct ggcggcggcg tcccacgact atgcccgcga gcagggtgg accggggccg 2280  
aggtccggcg cgaacgcggc atcctgccc ttgcgctggc ccatgacgcg gcgggcttct 2340  
gggccgatca cgcggagggg cctgttccc tgggactgcg cgcgggggtc tttcaccgg 2400  
tcaccggcta ttcgctgcc tatgcggcg aggtggcgga cgtggtggcg ggctgtccg 2460  
ggccgcccgg caccgacgcg ctgcgcggcg ccatccgga ttacgcgatc gaccgggcac 2520  
gccgtgaccg ctttctgcgc ctgctgaacc ggatgctgtt ccgcggctgc gcgccgcacc 2580  
ggcgctatac cctgctgcag cggttctacc gcatgccga tggactgatc gaacggttct 2640  
atgcccggcg gctgagcgtg gcggatcagc tgcgcctcgt gaccggcaag cctcccattc 2700  
cccttggcac ggccatccgc tgctgccc aacgtccct gctgaaggaa aacgcatgaa 2760  
cgccattcg cccgcggcca agaccgcat cgtgatcggc gcaggctttg gcgggctggc 2820  
cctggccatc cgcctgcagt ccgcgggcat cgccaccacc ctggtcgagg cccgggacaa 2880  
ggccggcggg cgcgcctatg tctggcacga tcagggccat gtcttcgacg cgggcccgc 2940  
cgtcatcacc gaccccgatg cgctcaagga gctgtggcg ctgaccgggc aggacatggc 3000  
gcgcgacgtg acgctgatgc cgggtgcgc cttctatcga ctgatgtggc cgggcgggaa 3060  
ggtcttcgat tacgtgaacg aggcgatca gctggagcgc cagatcgccc agttcaacct 3120  
ggacgacctg gaaggatacc gccgcttccg tgattacgcg gaggaggtgt atcaggaggg 3180  
ctacgtcaag ctgggcaccg tgcccttct caagctgggc cagatgctca aggcgcgcc 3240  
cgcgctgatg aagctggagg cctataagtc cgtccatgcc aaggtcgcga cttcatcaa 3300  
ggaccctat ctgcggcagg cgttttcgta tcacacgctg ctggtggcg ggaatccctt 3360

ctcgaccagc tcgatctatg cgctgatcca cgcgctggag cggcgcggcg gggctctggtt 3420  
 cgccaagggc ggcaccaacc agctggctgc gggcatggtc gcgctgttcg aacggcttgg 3480  
 cggccagatg atgctgaacg ccaaggtcgc ccggatcgag accgagggcg cgcggaccac 3540  
 gggcgtcacc ctggcggacg ggcggtcttt aagggccgac atggtcgcca gcaacggcga 3600  
 cgtcatgcac aactatcgcg acctgctggg ccacacggcc cgcgggcaga gccgcgcgaa 3660  
 atcgctggac cgcaagcgct ggtccatgtc gttgttcgtg ctgcatttcg gtctgcgcga 3720  
 ggcgccaag gacatcgcg atcacacccat cctgttcggc ccccgctaca gggagctggg 3780  
 caacgagatc ttcaagggcc cgaagctggc cgaggatttc tcgctgtacc tgcattcgcc 3840  
 ctgcacgacc gatccggaca tggcgcctcc gggcatgtcc acgcattacg tgctggcccc 3900  
 cgtgccgcat ctgggccgcg ccgagatcga ttgggcggtc gaggggccgc gctatgccga 3960  
 ccgcatacctg gcgtccctgg aggagcggct gatcccgaac ctgcgcgcca acctgaccac 4020  
 gacgcgcatac ttcacgcccg ccgatttcgc cagcgaactg aacgcccatac acggcagcgc 4080  
 cttctcggtc gagccgatcc tgacgcaatc cgcgtggttc cggccgcaca accgcgacaa 4140  
 gacgatccgc aacttctatc tggtcggcgc gggcacccat ccgggcgcgg gcattccggg 4200  
 cgtcgtgggc tcggccaagg ccacggccca ggtgatgctg tccgacctg cgggcgcatg 4260  
 agcgatctgg tcctgacctc gaccgaggcg atcacccaag ggtcgcaaag ctttgccacg 4320  
 gcggccaagc tgatgccgcc gggcatccgc gacgacacgg tgatgctcta tgcctggtgc 4380  
 cgccacgcgg atgacgtgat cgacggtcag gccctgggca gccgccccga ggcggtgaac 4440  
 gaccgcagc cgcggctgga cggcctgcgc gtcgacacgc tggcggccct gcagggcgac 4500  
 ggtccggtga cccgcctt tgccgcgctg cgcgcggtg cgcggcggca tgatttccc 4560  
 caggcctggc ccatggacct gatcgaaggc ttcgcgatgg atgtcgaggc gcgcgactat 4620  
 cgcacgctgg atgacgtgct ggaatattcc tatcacgtcg caggcatcgt cggcgtgatg 4680  
 atggcccgcg tgatgggcgt gcgcgacgat cctgtcctgg accgcgcctg cgacctgggg 4740  
 ctggcgttcc agctgaccaa catcgcgcgc gacgtgatcg acgatgcgcg catcgggcgg 4800  
 tgctatctgc cgggggactg gctggaccag gcgggcgcgc ggatcgacgg gccggtgccg 4860  
 tcgccggagc tgtacacagt gatcctccgg ctgttgatg aggcggaacc ctattacgcg 4920  
 tcggcgcggg tgggtctggc ggatctgcc cgcgctgcg cctgggtccat cgccgccgcg 4980  
 ctacggatct atcgcgccat cgggctgcgc atccgcaaga gcgggccgca ggcctatcgc 5040

cagcggatca gcacgtccaa ggctgccaa atcggcctgc tgggcgtcgg gggctgggat 5100  
gtcgcgcgat cagcctgcc gggggcgggc gtgtcgcggc agggcctctg gacccggccg 5160  
catcacgtct aggcgcgcgc ggcgtagggc agaaccggtt ccagcagggc cgcgatttcc 5220  
ggagcctgaa ggcgcttgct gcgcagcatc gcgtccagtt gggcgcggct ggcctcgtaa 5280  
tgacgggaca cgttctgcag gtctgacacg gccagaaggc cgcgccgcgg gccggggggc 5340  
gcggcatcgc gaccggtatc cttgccaagc gccgcctggt cgcggcggc gtccagcagg 5400  
tcgtcatagg actggaacac gcggcccagc tgacggccaa agtcgatcat ctgggtctgc 5460  
tcctcggcgt cgaactcctt gatcacggcc agcatctcca gcccggcgat gaacagcacg 5520  
ccggtcttca ggtcctgttc ctgttcgacc cccgcgccgt tcttgccgc gtgcaggctc 5580  
aggtcctggc cggcgcacag gccctgcggc cccagggacc gcgacaggat ccgcaccagc 5640  
tgcccccga ccgtgccga cgcgccgcgc gcaccggcca gcagggccat tgcctcggtg 5700  
atcagggcga tgccgccag cagggcacgg ctttcgccat gcgccacatg ggtcgcgggc 5760  
cggccgcggc gcagcccggc atcgtccatg cagggcaggc cgtcgaagat cagcgatgcg 5820  
gcatgcacca tctcgaccgc gcaggcggcg tcgacgatcg tgcgcagac cccgcccag 5880  
gcctctgccg caagcagcat cagcatgccg cggaaaccgc tgcccgcga cagcgcgcca 5940  
tggctcatgg ccgcgccgag cggctgcgac acggcaccga atccctgggc gatctcctca 6000  
agtctggtct gcagaagggt ggcgtggatc gggttgacgt ctgctctcat cagtgccttc 6060  
gcgcttggtt tctgacctgg cgggaaggtc agggcggggc ggcaccccg gaccgctcat 6120  
ccaccgtcaa cagtcccat gttggaacgg ttcacgcccg attgcgagcc ttttcgacgg 6180  
cgacgcgggg tcgcgcggca attgttcaa caaggtcagt gga 6223

<210> 5  
<211> 729  
<212> DNA  
<213> crtW gene

<400> 5  
atgagcgcac atgccctgcc caaggcagat ctgaccgcca ccagcctgat cgtctcgggc 60  
ggcatcatcg ccgcgtggct ggcctgcat gtgcatgcgc tgtggtttct ggacgcggcg 120  
gcgcatccca tcctggcgat cgcgaatttc ctggggctga cctggctgtc ggtcggctctg 180  
ttcttcatcg cgcgatgacg gatgcacggg tcggctcgtg cggggcgtcc gcgcggcaat 240  
gcggcgatgg gccagctggt cctgtggctg tatgccgat tttcgtggcg caagatgatc 300

gtcaagcaca tggcccatca ccgccataacc ggaaccgacg acgaccccga tttcgaccat 360  
ggcggcccg g tccgctggta cgcgcgcttc atcggcacct atttcggctg gcgcgagggg 420  
ctgctgctgc ccgtcatcgt gacggtctat gcgctgatcc tgggggatcg ctggatgtac 480  
gtggtcttct ggccgctgcc gtcgatacctg gcgtcgatcc agctgttcgt gttcggcacc 540  
tggctgccgc accgccccgg ccacgacgcg ttccccgacc gccataatgc gcggtcgtcg 600  
cggatcagcg accccgtgtc gctgctgacc tgctttcact ttggtgggta tcatcacgaa 660  
caccacctgc acccgacggt gccttggtgg cgcctgccca gcacccgcac caagggggac 720  
accgcatga 729

<210> 6  
<211> 242  
<212> PRT  
<213> crtW amino acid

<400> 6  
Met Ser Ala His Ala Leu Pro Lys Ala Asp Leu Thr Ala Thr Ser Leu  
1 5 10 15  
Ile Val Ser Gly Gly Ile Ile Ala Ala Trp Leu Ala Leu His Val His  
20 25 30  
Ala Leu Trp Phe Leu Asp Ala Ala Ala His Pro Ile Leu Ala Ile Ala  
35 40 45  
Asn Phe Leu Gly Leu Thr Trp Leu Ser Val Gly Leu Phe Phe Ile Ala  
50 55 60  
His Asp Ala Met His Gly Ser Val Val Pro Gly Arg Pro Arg Gly Asn  
65 70 75 80  
Ala Ala Met Gly Gln Leu Val Leu Trp Leu Tyr Ala Gly Phe Ser Trp  
85 90 95  
Arg Lys Met Ile Val Lys His Met Ala His His Arg His Thr Gly Thr  
100 105 110  
Asp Asp Asp Pro Asp Phe Asp His Gly Gly Pro Val Arg Trp Tyr Ala  
115 120 125  
Arg Phe Ile Gly Thr Tyr Phe Gly Trp Arg Glu Gly Leu Leu Leu Pro  
130 135 140  
Val Ile Val Thr Val Tyr Ala Leu Ile Leu Gly Asp Arg Trp Met Tyr  
145 150 155 160  
Val Val Phe Trp Pro Leu Pro Ser Ile Leu Ala Ser Ile Gln Leu Phe  
165 170 175  
Val Phe Gly Thr Trp Leu Pro His Arg Pro Gly His Asp Ala Phe Pro

	180		185		190
Asp Arg His Asn Ala Arg Ser Ser Arg Ile Ser Asp Pro Val Ser Leu					
	195		200		205
Leu Thr Cys Phe His Phe Gly Gly Tyr His His Glu His His Leu His					
	210		215		220
Pro Thr Val Pro Trp Trp Arg Leu Pro Ser Thr Arg Thr Lys Gly Asp					
	225		230		235
					240
Thr Ala					

<210> 7  
 <211> 489  
 <212> DNA  
 <213> crtZ gene

<400> 7  
 atgaccaatt tcctgatcgt cgtcgccacc gtgctggtga tggagttgac ggcctattcc 60  
 gtccaccggtt ggatcatgca cggccccctg ggctggggct ggcacaagtc ccaccacgag 120  
 gaacacgacc acgcgctgga aaagaacgac ctgtacggcc tggctctttgc ggtgatcgcc 180  
 acggtgctgt tcacggtggg ctggatctgg gcgccggtcc tgtggtggat cgctttgggc 240  
 atgaccgtct atgggctgat ctatttcgtc ctgcatgacg ggctggttca tcagcgctgg 300  
 ccgttccgct atatcccgcg caagggctat gcccgccgcc tgtatcaggc ccaccgcctg 360  
 caccacgcgg tcgagggacg cgaccattgc gtcagcttcg gcttcattca tgcgccgccg 420  
 gtcgacaagc tgaagcagga cctgaagacg tcgggctgctg tgcgggccga ggcgcaggag 480  
 cgcacgtga 489

<210> 8  
 <211> 162  
 <212> PRT  
 <213> crtZ amino acid

<400> 8  
 Met Thr Asn Phe Leu Ile Val Val Ala Thr Val Leu Val Met Glu Leu  
 1 5 10 15  
 Thr Ala Tyr Ser Val His Arg Trp Ile Met His Gly Pro Leu Gly Trp  
 20 25 30  
 Gly Trp His Lys Ser His His Glu Glu His Asp His Ala Leu Glu Lys  
 35 40 45  
 Asn Asp Leu Tyr Gly Leu Val Phe Ala Val Ile Ala Thr Val Leu Phe  
 50 55 60



Thr Val Gly Trp Ile Trp Ala Pro Val Leu Trp Trp Ile Ala Leu Gly  
 65 70 75 80  
 Met Thr Val Tyr Gly Leu Ile Tyr Phe Val Leu His Asp Gly Leu Val  
 85 90 95  
 His Gln Arg Trp Pro Phe Arg Tyr Ile Pro Arg Lys Gly Tyr Ala Arg  
 100 105 110  
 Arg Leu Tyr Gln Ala His Arg Leu His His Ala Val Glu Gly Arg Asp  
 115 120 125  
 His Cys Val Ser Phe Gly Phe Ile Tyr Ala Pro Pro Val Asp Lys Leu  
 130 135 140  
 Lys Gln Asp Leu Lys Thr Ser Gly Val Leu Arg Ala Glu Ala Gln Glu  
 145 150 155 160  
 Arg Thr

<210> 9  
 <211> 1161  
 <212> DNA  
 <213> crtY gene

<400> 9  
 gtgacccatg acgtgctgct ggcagggggcg ggccttgcca acgggctgat cgccctggcg 60  
 ctgcgcgcgg cgcgggcccga cctgcggggtg ctgctgctgg atcatgcggc gggaccgtca 120  
 gacggccata cctggctctg ccacgacccc gatctgtcgc cgcactggct ggcgcgggctg 180  
 aagcccctgc gccgcgcca ctggcccgc caggaggtgc gctttccccg ccatgcccgg 240  
 cggctggcca ccggttacgg gtcgctggac ggggcggcgc tggcggatgc ggtggcccgg 300  
 tcgggcgccc agatccgctg gaacagcgac atcgccctgc tggatgaaca gggggcgacg 360  
 ctgtcctgcg gcaccgggat cgaggcgggc gcggtcctgg acgggcgcgg cgcgagccg 420  
 tcgcggcatc tgaccgtggg tttccagaaa ttcgtgggcg tcgagatcga gaccgactgc 480  
 cccacggcg tgccccgcc gatgatcatg gacgcgaccg tcaccagca ggacgggtac 540  
 cgattcatct atctgctgcc cttctctccg acgcgcatcc tgatcgagga cactcgctat 600  
 tccgatggcg gcaatctgga cgacgacgcg ctggcggcgg cgtcccacga ctatgcccgc 660  
 cagcagggct ggaccggggc cgaggtccgg cgcgaaacgc gcatcctgcc cattgcgctg 720  
 gcccatgacg cggcgggctt ctgggccgat cacgcggagg ggctgttcc cgtgggactg 780  
 cgcgcggggt tctttcacc ggtcaccggc tattcgctgc cctatgcggc gcaggtggcg 840

gacgtggtgg cgggcctgtc cgggccgccc ggcaccgacg cgctgcgcgg cgccatccgc 900  
 gattacgcga tcgaccgggc acgccgtgac cgctttctgc gcctgctgaa ccggatgctg 960  
 ttccgcggct gcgcgcccga ccggcgctat accctgctgc agcggttcta ccgcatgccg 1020  
 catggactga tcgaacggtt ctatgccggc cggctgagcg tggcggatca gctgcgcac 1080  
 gtgaccggca agcctcccat tccccttggc acggccatcc gctgcctgcc cgaacgtccc 1140  
 ctgctgaagg aaaacgcatg a 1161

<210> 10  
 <211> 386  
 <212> PRT  
 <213> crtY amino acid

<400> 10  
 Val Thr His Asp Val Leu Leu Ala Gly Ala Gly Leu Ala Asn Gly Leu  
 1 5 10 15  
 Ile Ala Leu Ala Leu Arg Ala Ala Arg Pro Asp Leu Arg Val Leu Leu  
 20 25 30  
 Leu Asp His Ala Ala Gly Pro Ser Asp Gly His Thr Trp Ser Cys His  
 35 40 45  
 Asp Pro Asp Leu Ser Pro His Trp Leu Ala Arg Leu Lys Pro Leu Arg  
 50 55 60  
 Arg Ala Asn Trp Pro Asp Gln Glu Val Arg Phe Pro Arg His Ala Arg  
 65 70 75 80  
 Arg Leu Ala Thr Gly Tyr Gly Ser Leu Asp Gly Ala Ala Leu Ala Asp  
 85 90 95  
 Ala Val Ala Arg Ser Gly Ala Glu Ile Arg Trp Asn Ser Asp Ile Ala  
 100 105 110  
 Leu Leu Asp Glu Gln Gly Ala Thr Leu Ser Cys Gly Thr Arg Ile Glu  
 115 120 125  
 Ala Gly Ala Val Leu Asp Gly Arg Gly Ala Gln Pro Ser Arg His Leu  
 130 135 140  
 Thr Val Gly Phe Gln Lys Phe Val Gly Val Glu Ile Glu Thr Asp Cys  
 145 150 155 160  
 Pro His Gly Val Pro Arg Pro Met Ile Met Asp Ala Thr Val Thr Gln  
 165 170 175  
 Gln Asp Gly Tyr Arg Phe Ile Tyr Leu Leu Pro Phe Ser Pro Thr Arg  
 180 185 190  
 Ile Leu Ile Glu Asp Thr Arg Tyr Ser Asp Gly Gly Asn Leu Asp Asp  
 195 200 205

Asp Ala Leu Ala Ala Ala Ser His Asp Tyr Ala Arg Gln Gln Gly Trp  
 210 215 220  
 Thr Gly Ala Glu Val Arg Arg Glu Arg Gly Ile Leu Pro Ile Ala Leu  
 225 230 235 240  
 Ala His Asp Ala Ala Gly Phe Trp Ala Asp His Ala Glu Gly Pro Val  
 245 250 255  
 Pro Val Gly Leu Arg Ala Gly Phe Phe His Pro Val Thr Gly Tyr Ser  
 260 265 270  
 Leu Pro Tyr Ala Ala Gln Val Ala Asp Val Val Ala Gly Leu Ser Gly  
 275 280 285  
 Pro Pro Gly Thr Asp Ala Leu Arg Gly Ala Ile Arg Asp Tyr Ala Ile  
 290 295 300  
 Asp Arg Ala Arg Arg Asp Arg Phe Leu Arg Leu Leu Asn Arg Met Leu  
 305 310 315 320  
 Phe Arg Gly Cys Ala Pro Asp Arg Arg Tyr Thr Leu Leu Gln Arg Phe  
 325 330 335  
 Tyr Arg Met Pro His Gly Leu Ile Glu Arg Phe Tyr Ala Gly Arg Leu  
 340 345 350  
 Ser Val Ala Asp Gln Leu Arg Ile Val Thr Gly Lys Pro Pro Ile Pro  
 355 360 365  
 Leu Gly Thr Ala Ile Arg Cys Leu Pro Glu Arg Pro Leu Leu Lys Glu  
 370 375 380  
 Asn Ala  
 385

<210> 11  
 <211> 1506  
 <212> DNA  
 <213> crtI gene  
  
 <400> 11  
 atgaacgccc attcgcccg ccgaagacc gccatcgtga tcggcgcagg ctttggcggg 60  
 ctggccctgg ccatccgcct gcagtccgcg ggcatcgcca ccaccctggt cgaggcccg 120  
 gacaagcccc gcgggcgcg ctagtctctg cacgatcagg gccatgtctt cgacgcgggc 180  
 ccgaccgtca tcaccgaccc cgatgcgctc aaggagctgt gggcgctgac cgggcaggac 240  
 atggcgcgcg acgtgacgct gatgccggtg tcgcccttct atcgactgat gtggccgggc 300  
 gggaaggctc tcgattacgt gaacgaggcc gatcagctgg agcgccagat cgcccagttc 360  
 aaccgggacg acctggaagg ataccgccgc ttccgtgatt acgcggagga ggtgtatcag 420

```

gagggctacg tcaagctggg caccgtgccc ttcctcaagc tgggccagat gctcaaggcc 480
gcgcccgcgc tgatgaagct ggaggcctat aagtccgtcc atgccaaggt cgcgaccttc 540
atcaaggacc cctatctgcg gcaggcgttt tcgtatcaca cgctgctggt gggcggggaat 600
cccttctcga ccagctcgat ctatgcgctg atccacgcgc tggagcggcg cggcggggtc 660
tggttcgcca agggcggcac caaccagctg gtcgcgggca tggtcgcgct gttcgaacgg 720
cttggcggcc agatgatgct gaacgccaaag gtcgcccgga tcgagaccga gggcgcgcg 780
accacggggcg tcaccctggc ggacggggcg tctttaaggg ccgacatggt cgccagcaac 840
ggcgacgtca tgcacaacta tcgcgacctg ctgggccaca cggcccgcgg gcagagccgc 900
gcgaaatcgc tggaccgcaa gcgctggtcc atgtcgttgt tcgtgctgca tttcgggtctg 960
cgcgaggcgc ccaaggacat cgcgcatcac accatcctgt tcggcccccg ctacaggagg 1020
ctgggtcaacg agatcttcaa gggcccgaag ctggccgagg atttctcgct gtacctgcat 1080
tcgccctgca cgaccgatcc ggacatggcg cctccgggca tgtccacgca ttacgtgctg 1140
gccccgtgc cgcactctggg ccgcgccgag atcgattggg cggtcgaggg gccgcgctat 1200
gccgaccgca tcctggcgtc cctggaggag cggctgatcc cgaacctgcg cgccaacctg 1260
accacgacgc gcatcttcac gcccgccgat ttcgccagcg aactgaacgc ccatcacggc 1320
agcgccttct cggtcgagcc gatcctgacg caatccgcgt ggttcggcc gcacaaccgc 1380
gacaagacga tccgcaactt ctatctggtc ggcgcgggca cccatccggg cgcgggcatt 1440
ccgggcgtcg tgggctcggc caaggccacg gcccaggtga tgctgtccga cctggcgggc 1500
gcatga 1506

```

```

<210> 12
<211> 501
<212> PRT
<213> crtI amino acid

```

```

<400> 12
Met Asn Ala His Ser Pro Ala Ala Lys Thr Ala Ile Val Ile Gly Ala
 1             5             10             15

Gly Phe Gly Gly Leu Ala Leu Ala Ile Arg Leu Gln Ser Ala Gly Ile
      20             25             30

Ala Thr Thr Leu Val Glu Ala Arg Asp Lys Pro Gly Gly Arg Ala Tyr
      35             40             45

Val Trp His Asp Gln Gly His Val Phe Asp Ala Gly Pro Thr Val Ile
 50             55             60

```

Thr	Asp	Pro	Asp	Ala	Leu	Lys	Glu	Leu	Trp	Ala	Leu	Thr	Gly	Gln	Asp	
65					70					75					80	
Met	Ala	Arg	Asp	Val	Thr	Leu	Met	Pro	Val	Ser	Pro	Phe	Tyr	Arg	Leu	
				85					90					95		
Met	Trp	Pro	Gly	Gly	Lys	Val	Phe	Asp	Tyr	Val	Asn	Glu	Ala	Asp	Gln	
			100					105					110			
Leu	Glu	Arg	Gln	Ile	Ala	Gln	Phe	Asn	Pro	Asp	Asp	Leu	Glu	Gly	Tyr	
		115					120					125				
Arg	Arg	Phe	Arg	Asp	Tyr	Ala	Glu	Glu	Val	Tyr	Gln	Glu	Gly	Tyr	Val	
	130					135					140					
Lys	Leu	Gly	Thr	Val	Pro	Phe	Leu	Lys	Leu	Gly	Gln	Met	Leu	Lys	Ala	
145					150					155					160	
Ala	Pro	Ala	Leu	Met	Lys	Leu	Glu	Ala	Tyr	Lys	Ser	Val	His	Ala	Lys	
				165					170					175		
Val	Ala	Thr	Phe	Ile	Lys	Asp	Pro	Tyr	Leu	Arg	Gln	Ala	Phe	Ser	Tyr	
			180					185					190			
His	Thr	Leu	Leu	Val	Gly	Gly	Asn	Pro	Phe	Ser	Thr	Ser	Ser	Ile	Tyr	
		195					200					205				
Ala	Leu	Ile	His	Ala	Leu	Glu	Arg	Arg	Gly	Gly	Val	Trp	Phe	Ala	Lys	
	210					215					220					
Gly	Gly	Thr	Asn	Gln	Leu	Val	Ala	Gly	Met	Val	Ala	Leu	Phe	Glu	Arg	
225					230					235					240	
Leu	Gly	Gly	Gln	Met	Met	Leu	Asn	Ala	Lys	Val	Ala	Arg	Ile	Glu	Thr	
				245					250					255		
Glu	Gly	Ala	Arg	Thr	Thr	Gly	Val	Thr	Leu	Ala	Asp	Gly	Arg	Ser	Leu	
			260					265					270			
Arg	Ala	Asp	Met	Val	Ala	Ser	Asn	Gly	Asp	Val	Met	His	Asn	Tyr	Arg	
		275					280					285				
Asp	Leu	Leu	Gly	His	Thr	Ala	Arg	Gly	Gln	Ser	Arg	Ala	Lys	Ser	Leu	
	290					295					300					
Asp	Arg	Lys	Arg	Trp	Ser	Met	Ser	Leu	Phe	Val	Leu	His	Phe	Gly	Leu	
305					310					315					320	
Arg	Glu	Ala	Pro	Lys	Asp	Ile	Ala	His	His	Thr	Ile	Leu	Phe	Gly	Pro	
				325					330					335		
Arg	Tyr	Arg	Glu	Leu	Val	Asn	Glu	Ile	Phe	Lys	Gly	Pro	Lys	Leu	Ala	
			340					345					350			
Glu	Asp	Phe	Ser	Leu	Tyr	Leu	His	Ser	Pro	Cys	Thr	Thr	Asp	Pro	Asp	
		355					360					365				

Met Ala Pro Pro Gly Met Ser Thr His Tyr Val Leu Ala Pro Val Pro  
 370 375 380  
 His Leu Gly Arg Ala Glu Ile Asp Trp Ala Val Glu Gly Pro Arg Tyr  
 385 390 395 400  
 Ala Asp Arg Ile Leu Ala Ser Leu Glu Glu Arg Leu Ile Pro Asn Leu  
 405 410 415  
 Arg Ala Asn Leu Thr Thr Thr Arg Ile Phe Thr Pro Ala Asp Phe Ala  
 420 425 430  
 Ser Glu Leu Asn Ala His His Gly Ser Ala Phe Ser Val Glu Pro Ile  
 435 440 445  
 Leu Thr Gln Ser Ala Trp Phe Arg Pro His Asn Arg Asp Lys Thr Ile  
 450 455 460  
 Arg Asn Phe Tyr Leu Val Gly Ala Gly Thr His Pro Gly Ala Gly Ile  
 465 470 475 480  
 Pro Gly Val Val Gly Ser Ala Lys Ala Thr Ala Gln Val Met Leu Ser  
 485 490 495  
 Asp Leu Ala Gly Ala  
 500

<210> 13  
 <211> 915  
 <212> DNA  
 <213> crtB gene

<400> 13  
 atgagcgcac tggtcctgac ctcgaccgag gcgatcaccc aagggtcgca aagctttgcc 60  
 acggcgccca agctgatgcc gccgggcatc cgcgacgaca cggtgatgct ctatgcctgg 120  
 tgccgccacg cggatgacgt gatcgacggt caggccctgg gcagccgccc cgaggcgggtg 180  
 aacgacccgc aggcgcggct ggacggcctg cgcgtcgaca cgctggcggc cctgcagggc 240  
 gacgggtccg tgaccccgcc ctttgccgcg ctgcgcgcgg tggcgcggcg gcatgatttc 300  
 ccgcaggcct ggcccatgga cctgatcgaa ggcttcgcga tggatgtcga ggcgcgcgac 360  
 tatcgacgcg tggatgacgt gctggaatat tcctatcacg tcgcaggcat cgtcggcgtg 420  
 atgatggccc gcgtgatggg cgtgcgcgac gatcctgtcc tggaccgcgc ctgcgacctg 480  
 gggctggcgt tccagctgac caacatcgcg cgcgacgtga tcgacgatgc gcgcacgcgg 540  
 cgggtgctatc tgccggggga ctggctggac caggcggggc cgcggatcga cgggcccggg 600  
 ccgtcgccgg agctgtacac agtgatcctc cggctgttgg atgaggcgga accctattac 660

gcgtcggcgc ggggtgggtct ggcggatctg ccaccgcgct gcgcctggtc catcgccgcc 720  
 gcgctacgga tctatcgcgc catcgggctg cgcatccgca agagcgggcc gcaggcctat 780  
 cgccagcgga tcagcacgtc caaggctgcc aagatcggcc tgctgggcgt cgggggctgg 840  
 gatgtcgcgc gatcacgcct gccgggggcg ggcgtgtcgc ggcagggcct ctggaccgg 900  
 ccgcatcacg tctag 915

<210> 14  
 <211> 304  
 <212> PRT  
 <213> crtB amino acid

<400> 14  
 Met Ser Asp Leu Val Leu Thr Ser Thr Glu Ala Ile Thr Gln Gly Ser  
 1 5 10 15  
 Gln Ser Phe Ala Thr Ala Ala Lys Leu Met Pro Pro Gly Ile Arg Asp  
 20 25 30  
 Asp Thr Val Met Leu Tyr Ala Trp Cys Arg His Ala Asp Asp Val Ile  
 35 40 45  
 Asp Gly Gln Ala Leu Gly Ser Arg Pro Glu Ala Val Asn Asp Pro Gln  
 50 55 60  
 Ala Arg Leu Asp Gly Leu Arg Val Asp Thr Leu Ala Ala Leu Gln Gly  
 65 70 75 80  
 Asp Gly Pro Val Thr Pro Pro Phe Ala Ala Leu Arg Ala Val Ala Arg  
 85 90 95  
 Arg His Asp Phe Pro Gln Ala Trp Pro Met Asp Leu Ile Glu Gly Phe  
 100 105 110  
 Ala Met Asp Val Glu Ala Arg Asp Tyr Arg Thr Leu Asp Asp Val Leu  
 115 120 125  
 Glu Tyr Ser Tyr His Val Ala Gly Ile Val Gly Val Met Met Ala Arg  
 130 135 140  
 Val Met Gly Val Arg Asp Asp Pro Val Leu Asp Arg Ala Cys Asp Leu  
 145 150 155 160  
 Gly Leu Ala Phe Gln Leu Thr Asn Ile Ala Arg Asp Val Ile Asp Asp  
 165 170 175  
 Ala Arg Ile Gly Arg Cys Tyr Leu Pro Gly Asp Trp Leu Asp Gln Ala  
 180 185 190  
 Gly Ala Arg Ile Asp Gly Pro Val Pro Ser Pro Glu Leu Tyr Thr Val  
 195 200 205

Ile Leu Arg Leu Leu Asp Glu Ala Glu Pro Tyr Tyr Ala Ser Ala Arg  
 210 215 220  
 Val Gly Leu Ala Asp Leu Pro Pro Arg Cys Ala Trp Ser Ile Ala Ala  
 225 230 235 240  
 Ala Leu Arg Ile Tyr Arg Ala Ile Gly Leu Arg Ile Arg Lys Ser Gly  
 245 250 255  
 Pro Gln Ala Tyr Arg Gln Arg Ile Ser Thr Ser Lys Ala Ala Lys Ile  
 260 265 270  
 Gly Leu Leu Gly Val Gly Gly Trp Asp Val Ala Arg Ser Arg Leu Pro  
 275 280 285  
 Gly Ala Gly Val Ser Arg Gln Gly Leu Trp Thr Arg Pro His His Val  
 290 295 300

<210> 15  
 <211> 882  
 <212> DNA  
 <213> crtE gene

<400> 15  
 atgagacgag acgtcaaccc gatccacgcc acccttctgc agaccagact tgaggagatc 60  
 gccagggat tcggtgccgt gtcgcagccg ctccggcgcg ccatgagcca tggcgcgctg 120  
 tcgtcgggca ggcggttccg cggcatgctg atgctgcttg cggcagaggc ctccggcggg 180  
 gtctgcgaca cgatcgtcga cgccgcctgc gcggtcgaga tggatcatgc cgcacgctg 240  
 atcttcgacg acctgcctg catggacgat gccgggctgc gccgcggccg gcccgcgacc 300  
 catgtggcgc atggcgaaag ccgtgccgtg ctggggcgga tcgccctgat caccgaggca 360  
 atggccctgc tggccggtgc gcgcggcgcg tcgggcacgg tgcgggcgca gctgggtcgg 420  
 atcctgtcgc ggtccctggg gccgcagggc ctgtgcgccg gccaggacct ggacctgcac 480  
 gcggccaaga acggcgcggg ggtcgaacag gaacaggacc tgaagaccgg cgtgctgttc 540  
 atcgccgggc tggagatgct ggccgtgatc aaggagtctg acgccgagga gcagaccag 600  
 atgatcgact ttggccgtca gctggggcgc gtgttccagt cctatgacga cctgctggac 660  
 gtcgtgggcg accaggcggc gcttggcaag gataccggtc gcgatgccgc ggccccggc 720  
 ccgcggcgcg gccttctggc cgtgtcagac ctgcagaacg tgtcccgtca ttacgaggcc 780  
 agccgcgccc aactggacgc gatgctgcgc agcaagcgcc ttcaggctcc ggaaatcgcg 840  
 gccctgctgg aacgggttct gccctacgcc gcgcgcgcct ag 882



<210> 16  
 <211> 293  
 <212> PRT  
 <213> crtE amino acid

<400> 16

Met	Arg	Arg	Asp	Val	Asn	Pro	Ile	His	Ala	Thr	Leu	Leu	Gln	Thr	Arg
1				5					10					15	
Leu	Glu	Glu	Ile	Ala	Gln	Gly	Phe	Gly	Ala	Val	Ser	Gln	Pro	Leu	Gly
			20					25					30		
Ala	Ala	Met	Ser	His	Gly	Ala	Leu	Ser	Ser	Gly	Arg	Arg	Phe	Arg	Gly
		35					40					45			
Met	Leu	Met	Leu	Leu	Ala	Ala	Glu	Ala	Ser	Gly	Gly	Val	Cys	Asp	Thr
	50					55					60				
Ile	Val	Asp	Ala	Ala	Cys	Ala	Val	Glu	Met	Val	His	Ala	Ala	Ser	Leu
65					70				75						80
Ile	Phe	Asp	Asp	Leu	Pro	Cys	Met	Asp	Asp	Ala	Gly	Leu	Arg	Arg	Gly
				85					90					95	
Arg	Pro	Ala	Thr	His	Val	Ala	His	Gly	Glu	Ser	Arg	Ala	Val	Leu	Gly
			100					105					110		
Gly	Ile	Ala	Leu	Ile	Thr	Glu	Ala	Met	Ala	Leu	Leu	Ala	Gly	Ala	Arg
	115						120					125			
Gly	Ala	Ser	Gly	Thr	Val	Arg	Ala	Gln	Leu	Val	Arg	Ile	Leu	Ser	Arg
	130					135					140				
Ser	Leu	Gly	Pro	Gln	Gly	Leu	Cys	Ala	Gly	Gln	Asp	Leu	Asp	Leu	His
145					150					155					160
Ala	Ala	Lys	Asn	Gly	Ala	Gly	Val	Glu	Gln	Glu	Gln	Asp	Leu	Lys	Thr
				165					170					175	
Gly	Val	Leu	Phe	Ile	Ala	Gly	Leu	Glu	Met	Leu	Ala	Val	Ile	Lys	Glu
		180					185						190		
Phe	Asp	Ala	Glu	Glu	Gln	Thr	Gln	Met	Ile	Asp	Phe	Gly	Arg	Gln	Leu
		195					200					205			
Gly	Arg	Val	Phe	Gln	Ser	Tyr	Asp	Asp	Leu	Leu	Asp	Val	Val	Gly	Asp
	210					215					220				
Gln	Ala	Ala	Leu	Gly	Lys	Asp	Thr	Gly	Arg	Asp	Ala	Ala	Ala	Pro	Gly
225					230					235					240
Pro	Arg	Arg	Gly	Leu	Leu	Ala	Val	Ser	Asp	Leu	Gln	Asn	Val	Ser	Arg
				245					250					255	
His	Tyr	Glu	Ala	Ser	Arg	Ala	Gln	Leu	Asp	Ala	Met	Leu	Arg	Ser	Lys
		260					265						270		

Arg Leu Gln Ala Pro Glu Ile Ala Ala Leu Leu Glu Arg Val Leu Pro  
 275 280 285

Tyr Ala Ala Arg Ala  
 290

<210> 17  
 <211> 19  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> forward primer for crt gene

<400> 17  
 gttccacgac tggggcatc 19

<210> 18  
 <211> 28  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> reverse primer for crt gene

<400> 18  
 tccactgacc ttgttgaca aattgccg 28